An Integrated Approach for Analysing and Assessing the Performance of Virtual Learning Groups

Thanasis Daradoumis[1], Alejandra Martínez-Monés[2], and Fatos Xhafa[1]

[1] Open University of Catalonia, Department of Information Sciences
Av. Tibidabo, 39-43, 08035 Barcelona, Spain
{adaradoumis,fxhafa}@uoc.edu
[2] University of Valladolid, Department of Computer Science
ETIT, Campus Miguel Delibes, 47011 Valladolid, Spain
amartine@infor.uva.es

Abstract. Collaborative distance learning involves a variety of elements and factors that have to be considered and measured in order to analyse and assess group and individual performance more effectively and objectively. This paper presents an approach that integrates qualitative, social network analysis (SNA) and quantitative techniques for evaluating online collaborative learning interactions. Integration of various different data sources, tools and techniques provides a more complete and robust framework for group modelling and guarantees a more efficient evaluation of group effectiveness and individual competence. Our research relies on the analysis of a real, long-term, complex collaborative experience, which is initially evaluated in terms of principled criteria and a basic qualitative process. At the end of the experience, the coded student interactions are further analysed through the SNA technique to assess participatory aspects, identify the most effective groups and the most prominent actors. Finally, the approach is contrasted and completed through a statistical technique which sheds more light on the results obtained that far. The proposal draws a well-founded line toward the development of a principled framework for the monitoring and analysis of group interaction and group scaffolding which can be considered a major issue towards the actual application of the CSCL proposals to real classrooms.

1 Introduction

The application of innovative educational experiences proposed by research fields such as CSCL is hindered by the obstacles that teachers find in applying them in real classrooms. For example, it is normal that students find difficulties while fulfilling the collaborative assignments and it is nearly impossible for a single teacher to account for all the interaction breakdowns that may happen within a medium-sized classroom. Moreover, if the students' collaborative process is to be taken into account for their assessment, teachers need efficient and effective tools that help them to perform this task. Therefore, there is a need of new approaches and tools that assist teachers in providing formative evaluation and assessment on collaboration in order to be able to apply these techniques to real classrooms.

These questions are closely related to the extensive research on interaction analysis in the CSCW and CSCL fields that has been oriented to identifying and exploring the

factors that affect the effectiveness and success of online group work and learning [8]. However, this line of research has proposed rather limited approaches, focusing on a single collaboration channel, such as dialogue [1] or action [13]. Some researchers have recently proposed an integration of different sources of data in the analysis [4, 15]. In spite of this fact, existing approaches have not yet managed to meet these needs satisfactorily, since most of them focus on experimental situations, which do not exactly reflect the issues and problems of a real situation.

Teachers need principled frameworks that help them to carry out all these issues. The evaluation of collaborative learning has to be performed at least at two levels, separating the process (or group functioning) from the product (or task performance) of collaboration [6, 10]. Evaluation of task performance considers the task skills and knowledge acquired by each member as well as the quality of final product of group work. Evaluation of group functioning refers to the analysis and assessment both of the interaction behaviour of group members and the social aspects of group work. According to [14], participation is a further important aspect since, together with acquisition, constitutes one of the two main metaphors of learning.

The complexity of the interaction processes and the need of considering different perspectives claim for *mixed evaluation methods* [9] that integrate different sources of data (i.e. types of interaction, products, student's opinions, etc.) and different analysis techniques in order to tackle with all the points of view that must be considered. To our knowledge, there are still very few mixed evaluation approaches that combine different data sources and methods to provide a more in-depth analysis of collaborative learning interactions.

We have been working in defining a principled, effective and holistic framework for analysing the interaction and assessing the performance of virtual learning groups [7]. This framework can be considered as a first step towards the definition of an evaluation approach that considers both the process and the product of collaboration. However, it still lacked some aspects, mainly related to the participatory metaphor, such as the general structure of the collaborative relationships and the roles that the students play with respect to these structures.

Recently, [12] has proposed a *mixed method* for the formative evaluation of social aspects of CSCL experiences that integrates different data sources and methods of analysis, including SNA [16], in order to evaluate social aspects of group work. SNA seeks to describe patterns of relationships among actors, to analyse the structure of these patterns and discover their effects on people and organizations. Social networks can be visualized as graphs called *sociograms*, which represent the actors as nodes of the graphs and the links among them as lines in the graph. Several studies have demonstrated its value within the CSCL field for the study of structural properties of individuals learning in groups [3, 18]. However, SNA by itself is not enough for achieving a full understanding of the collaborative processes. It needs to be complemented with other methods and perspectives.

The work we present in this paper can be considered as a new step towards the fulfilment of the demands of teachers for the formative evaluation in CSCL real settings, by integrating these proposals in a common evaluation experience. It shows how to use qualitative, quantitative and SNA techniques to analyse and assess a real case study.

The rest of the paper is structured as follows. First, we present the case study on which the analysis was carried out. On the one hand, the proposed approach is justi-

fied, the evaluation criteria (indicators) are set and a basic qualitative evaluation process is defined together with the first level of the mixed evaluation scheme we are proposing. On the other hand, in order to accomplish all evaluation criteria, two further levels of analysis are defined, including a quantitative and SNA evaluation technique (as well as the associated data sources and supporting tools). Section 3 describes the SNA and quantitative techniques in more detail and draws the final conclusions that result from the combination of all the three evaluation techniques. Finally, we present the future lines of our research that aim at building a principled framework for group modelling and scaffolding.

2   Case Study Description and Evaluation Method Proposed

The context of our research is defined by setting virtual learning groups to work on a real, long-term, complex, collaborative problem-solving situation that forms part of a distance learning undergraduate course. As such, it is important to present first a sufficient description of the workings of the case study that was used for the purpose of our analysis. Then we proceed to justify our approach and set the qualitative basis of the methodology proposed to tackle with the complex task of analysing and assessing group interaction successfully. Finally, the rest of the data sources, evaluation techniques and supporting tools that compose our approach are explained.

2.1   Case Study Description

To perform this study we were based on a real collaborative learning experience that was carried out in the scope of a virtual (distance) learning undergraduate interdisciplinary course, called "Application of Information Systems to Business". The experience run over a period of 14 weeks and involved 2 tutors and 122 students distributed into 21 virtual groups of 5 to 6 members. Students had to collaborate and work out a case study that simulated a real project in a business or organisation.

Virtual groups were formed and consolidated during the first 10 days of the course by the students themselves, following a well-structured and guided virtual process, called *group formation process*, based on the work of [5]. The case resolution consists of a set of target goals (phases) that are realised collaboratively (except the first one which aims at studying and understanding the case presented).

The whole project was carried out mostly asynchronously; synchronous interaction occurred in few specific cases of decision-making. All asynchronous collaborative interactions took place on the Basic Support for Cooperative Work (BSCW) system, a groupware tool that enables asynchronous and synchronous collaboration over the web [2]. BSCW offers shared workspaces that groups can use to store, manage, jointly edit and share documents, realise threaded discussions, etc.

To structure the whole collaborative learning process, we set two particularised shared workspaces in the BSCW system. The first one is a general workspace, which can be accessed by all students of the virtual class. The main purpose of this workspace is to let the students interact with each other in order to form the virtual learning groups. In addition, it is used to effectuate specific debates, which form part of the project requirements and involve all students, as well as to share important informa-

tion about the project among tutors and students. The other workspace type is a private space designated to house each virtual group, that is to record and structure the interaction of its members that aims to achieve the project target goals through the resolution of the specific tasks and problems the project consists of.

Table 1. Description of the evaluation criteria defined for each one of the aspects of the collaborative learning analysis. For each aspect we give the corresponding weight and the weights of its criteria.

Evaluation Criteria		Weight
Task performance		**50%**
TP1	The students' individual and group problem-solving capabilities and learning outcomes (*acquisition* metaphor)	40%
TP2	The students' contributing behaviour during task realisation (production function and use of active learning skills)	40%
TP3	The students' individual and group ongoing (and final) performance in terms of self-evaluation	20%
Group functioning		**20%**
GF1	Active participation behaviour	30%.
GF2	Social grounding (well-balanced contributions and role playing)	20%
GF3	Active interaction or processing skills that monitor and facilitate the group's well-being function	30%
GF4	Group processing (examine whether each member learnt how to interact and collaborate more effectively with his/her teammates)	20%
Social support		**15%**
SS1	Members' commitment toward collaboration, joint learning and accomplishment of the common group goal	30%
SS2	Level of peer involvement and their influential contribution to the involvement of the others	30%
SS3	Members' contribution to the achievement of mutual trust	10%
SS4	Members' motivational and emotional support to their peers	20%
SS5	Participation and contribution to conflict resolution	10%
Help services		**15%**
HS1	Help is timely	25%
HS2	Help is relevant to the student's needs	10%
HS3	Help is qualitative	30%
HS4	Help is understood by the student	25%
HS5	Help can be readily applied by the student	10%

Our analysis was carried out at both the general and the private group spaces, using specific evaluation criteria as parameters to measure the groups' real effectiveness regarding learning and collaborative skills, as explained in the following sections.

2.2 Defining Principled Evaluation Criteria and a Qualitative Evaluation Process

The need for a mixed evaluation scheme comes forth from identifying the most important indicators related to group activity which, due to its variety, can not be covered and satisfied by a unique analysis method. The first step toward establishing a

well suited and effective integrated evaluation approach is identifying principled evaluation criteria or group activity indicators (and their estimated weights) that capture and describe group interaction and performance sufficiently in the context where group work and learning is situated. The second step consists in presenting different analysis techniques to meet the evaluation criteria proposed and testing that neither technique alone nor a combination of any subset of them can account for all the indicators, thus prompting the necessity of an integrated approach. Consequently, the weight (and importance) of each technique within the integrated approach is estimated in terms of the indicators that accomplishes. Next we proceed to define the criteria and the first level of our approach, a basic qualitative evaluation process.

Based on the theoretical principles and indicators of effective collaboration of [10, 11, 14,15,17] we specify four important levels or aspects of collaborative learning analysis: *task performance* (or learning outcome), *group functioning* (or participation/ interaction behaviour), *social support*, and *help supply* (or task/process scaffolding); for more details on these levels see [6].

To measure and evaluate each level, we need to define generic evaluation criteria that describe and capture its important features as fully as possible. Each criterion is also assigned a specific weight. This is an important feature of our approach since it determines not only the importance of each evaluation means but also the way these means can be combined to carry out the analysis and evaluation process. The criteria identification is based on the above theoretical principles and our lengthy experience with online collaborative learning teams; the latter is also an influencing factor for specifying the weights of each criterion. Moreover, particularisation of the evaluation criteria and their associated weights depends on premises, such as the evaluation goals, the context or situation surrounding the collaborative learning experience and its specific tasks, as well as the available evaluation techniques and data sources.

In general, what we describe below shows a way to set weights for the analysis and assessment of our particular case study. We are currently exploring a more principled mechanism (such as a regression statistic model) to derive relative weights for each indicator. Table 1 presents both the evaluation criteria and the assigned weights.

In this experience we consider task performance an important factor of the student evaluation; in addition, we can analyse it through a variety of data sources and methods, thus we assign it a 50% weight. Among the other three factors, we consider that, at the time being, group functioning can be measured more easily and effectively than social support and help supply. Taking this condition into account, we assign it a 20% weight, whereas the other two factors are assigned an equal 15% weight each.

Our case study offers the tutor the context to perform a continuous qualitative evaluation of the students' work and collaborative activity. Thus, a formative qualitative evaluation takes up an important value and constitutes the basis of the evaluation method. In fact, all the four analysis aspects are measured and assessed qualitatively by the tutor at the end of each project phase.

In particular, as concerns task performance analysis, each group delivers the tutor its learning outcomes (the solution of the current sub-problem). The tutor corrects and assesses it thoroughly, assigns it a mark and sends his/her feedback to the group (criterion TP1). In addition, during task realisation the tutor performs a selective qualitative examination of the students' most significant contributions to the task, with the aim to reason out both the specific production function and the active learning skills exhibited by each group member (criterion TP2). Finally, the tutor carries out a quali-

tative analysis and assessment of two self-evaluation report types delivered by the students. The first one is elaborated by each group at the end of each problem-solving phase, whereas the latter is addressed to each student individually at the end of the project. Both reports are guided by specific questions that aim at knowing the students' personal opinion, perception and impression about individual contribution and overall group performance regarding the task (criterion TP3).

Such a qualitative evaluation is not only important to assess task performance issues but also to explore and qualify aspects related to the other three analysis levels (group functioning, social support and help supply). Individual and group self-evaluation reports proved to be a valuable information source for the tutor to delineate and get a basic comprehension of the groups' internal workings such as to explore: the members' participation trends and interaction or processing skills (criteria GF1, GF3), the achievement of a well-balanced group and adequate role playing (criterion GF2), the members' ability to reflect upon the way they learn to collaborate with each other (criterion GF4), the level of group cohesion accomplished (criteria SS1 to SS5), and the degree and suitability of the help provided by each member (criteria HS1 to HS5).

As a result, at this stage of our work, given that the qualitative evaluation is spread along all analysis levels, it is considered as a basic initial layer upon which further evaluation techniques can be incorporated and applied so that to fill the gaps left and evaluate all those issues that could not be covered sufficiently by the qualitative evaluation method alone, thus giving rise to a mixed evaluation approach.

2.3 Further Evaluation Techniques, Data Sources and Supporting Tools

Evidently the above qualitative evaluation method alone does not suffice to provide a complete analysis and assessment of all the indicators we identified at the four analysis levels. Indeed, self-evaluation reports (as well as further possible interviews or questionnaires) can give important information to the tutor, especially as regards issues related to social support and help provided by group members, which otherwise is difficult to extract from other sources or methods. However, there are still some important indicators, which need a further in-depth analysis in order to be evaluated completely. Such indicators are: TP2, GF1, GF2 and GF3.

The nature and objective of these four indicators suggests us to group them into two categories. The first one comprises criteria TP2 and GF3 that intend to measure specific students' active skills employed either to support task realisation and learning or to facilitate better interaction respectively. The second category involves criteria GF1 and GF2 that study participatory issues (like active peer involvement) and social grounding aspects of learning (like comparison of group activity, well-balanced interaction and role playing) respectively.

On the one hand, SNA has proved to be an adequate and sufficient technique to analyse the structure of the social interactions that take place in the virtual workspaces (criteria GF1 and GF2). This structure allows for the study of individual properties (prominence of the actors), small groups and the whole network.

On the other hand, to account for and evaluate the many different types of students' active learning or interaction/processing skills (criteria TP2 and GF3), we proceed to measure (quantify) the types of student contributions that imply a particular

skill; to that end, due to the big amount of interaction data produced, a quantitative (statistical) technique can provide better and more reliable results.

The basic source that provides data for both types of analysis is the BSCW *daily log files*. Every log file records all the interaction data (events) occurred in all active BSCW workspaces. BSCW distinguishes and generates four generic types of events (or actions) related to an object: *Create*, *Change*, *Read* and *Move* events.

In this study we take three major objects into account: *folder* (related to task and knowledge management), *document* (related to task realisation and learning), and *note(s)* (related to communication processing). The rest, which are considered minor objects, are grouped into a class called *others* and include, among others, concepts such as URL, Appointment, User, or Group Agenda.

On the one hand, to facilitate the quantitative analysis of the many different action types a student can perform, we initially use a specific software tool that extracts and filters the data contained in the event logs according to desired parameters defined by our analysis needs (for instance, events can be classified by user and action type, or can be distributed in specific periods of time).

On the other hand, to support SNA automatic processing, we use a tool called SAMSA (System for Adjacency Matrix and Sociogram-based Analysis). This tool contains several input modules, one of which takes data from BSCW event logs and transforms them into an XML file representing the interactions. Then, SAMSA allows us to select and configure the network we want to study (selecting dates, actors, and relationship type). The tool builds the matrix that represents the network, known as *sociomatrix*, and computes the indexes chosen for the analysis. It also shows the sociogram representing the network, and allows for visualising the actors' attributes.

Both analysis techniques are discussed in the next section and confirm that if the qualitative method were used alone, the students' evaluation would be insufficient and subjective. Instead, the combination of all three techniques together provides a more effective, complete and correct evaluation. This is particularly shown by the results obtained from the analysis of a specific effective group.

3 Analysing Group Interactions and Social Aspects of Group Work

This section describes the analysis results obtained by the SNA and descriptive statistical quantitative techniques. Each technique aims at achieving specific evaluation criteria set in the previous section and thus is complementary to the other. The obtained results are finally compared to the qualitative evaluation results and definite conclusions are drawn with respect to the real effectiveness of the learning group that was analysed and evaluated as an example. First the SNA technique is described, showing how SNA can be used in a top-down analysis approach so that the evaluator can start from a very general perspective of the classroom interactions and detect which are the groups or actors that need further analysis.

3.1   Analysis of the Participatory Aspects
     of the Collaborative Learning Processes by Social Network Analysis

In order to perform Social Network Analysis, it is necessary to define the networks and the set of indexes that will be used for the study.

Networks are *relationships* established among a set of *actors*. In this study we considered the relatioships composed by the *indirect links* between an actor that creates an object in the BSCW workspace and those that access this object in order to read it. This is by far the most frequent type of interaction in the context of the use of BSCW as shown by the daily report and the log files maintained by the system. The set of actors included both the students and the teacher or teachers.

We have identified a set of SNA indicators for the study of participatory aspects of learning, which were the ones used in this study, namely: *Network density* (*D*), *actor's degree centrality* ($C_D(n_i)$), and *network degree centralization* (C_D) [16]. *D* measures how knitted a network is, with values ranging from 0 (most sparse) to 1 (most dense). *Degree centrality* is an index of the actor's prestige. Given an actor n_i, $C_D(n_i)$ is the proportion of actors that are adjacent to n_i. It reflects the activity of the actors. In the case of directed relationships that consider the direction of the link, two degree indexes are defined: *indegree*, or the number of links terminating at the node; and *outdegree*, or the number of links originating at the node. Finally, *network degree centralization* (C_D), is a group-level measure based on actor's degree centrality. It gives an idea about the dependency of the network on the activity of a small group of actors. Its values range from 0 (even distribution of activity) to 1 (most centralized network). Directed networks define the corresponding indexes of *indegree centralization* (C_{ID}) and *outdegree centralization* (C_{OD}). All of these indexes and ranges apply to dichotomous relationships that can have only one out of two possible values: 0 when there is no link and 1 when there is a link between two actors. It is also possible to consider valued relationships that include a number showing their strength. The indexes computed on these relationships are more difficult to generalize than those computed from the dichotomous relationships, but sometimes are important to provide additional information. All of these indexes provide basic information about the activity of the actors in the network and about the global structure of the network according to different relationships. Moreover, they are simple to understand and to interpret, which are important features for facilitating their use by evaluators, who are not expected to be experts in SNA methods.

These indexes were applied in order to study and compare interactions at the general and at the private workspaces, as well as to identify who were the more and the less active students at both levels. The following two subsections develop these issues.

3.1.1   Student Participation in the General Workspace

The first aspect we wanted to analyse was the general structure of the relationships in the classrooms, which was studied by the indirect relationships network at the general workspace of the virtual classroom.

A first analysis can be performed by studying the network of the aggregated relationships during the last four phases of the course, once the groups had been created

and the students were focused on their project-oriented tasks. The indexes of this network ($\Delta=14,24\%$, $C_{ID}=42,22\%$, $C_{OD=}63,33\%$) show that the indirect links considered in this network were quite frequent (if we take into account the size of the network) but too much centralised as regards both reading (C_{ID}) and specially writing (C_{OD}), which means that the activity was concentrated on a very reduced set of actors.

The examination of the sociogram of this network (Fig. 1) allows us to go deeper in this issue in a very intuitive manner that can be easily built by the teacher or evaluator by means of SAMSA. The actors are represented by different shapes according to the group they belong to, and the links between them as directed lines that go from the actor that creates a document to the one(s) that read it. While the high concentration of the lines and arrows makes difficult a detailed analysis of the specific links, it is still possible to draw some initial conclusions that complement the values mentioned above. Firstly, it can be observed that some actors appear as isolated nodes at the left, which means that they did not intervene in the shared workspace at all. It is also possible to see that the teacher (*ifa*) has a central position in the network, shared with an important number of students. It is possible to identify at a glance who were the most active students (at the centre) and the less active ones (at the periphery). For example, actors like *xva, cao, dac,* and *fca* played a prominent role, while others, like *jga, Ton, ngu, etc.* played a peripheral role, as their only connection to it consists of a single link to another actor.
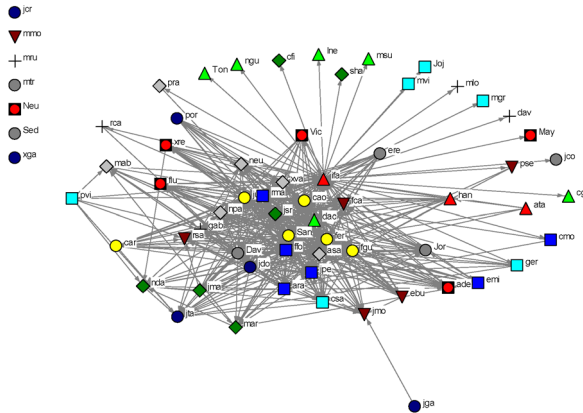

Fig. 1. Indirect relationships network at the general workspace during the course.

The exploration of the actors' centrality values, which is also calculated by SAMSA, complements the analysis of the sociogram, as they allow identifying the most active students, with the added value that these indexes inform also about the reading and writing activities. It is especially interesting to detect who were the students with a higher value of out-degree centralization, which means that they wrote documents that were read by more students. According to [3], this index can be considered as a measure of actors' prestige, and can help a teacher to detect who are the students which act as leaders in the sense measured by the network. In our study the students with a higher out-degree centrality value were *san, fer, jur, fgu,* and *car*.

The analyisis performed so far shows a static view of the aggregated relationships during the course. This view was complemented with an analysis of the evolution of the networks, which is performed by a similar study of each the phases in which the course was divided. It is not possible to show the details of this evolution for lack of space, but we comment here the main results: density remained stable throught the course (with values around 5-6% in each one of the phases), with a slight drop at the last phase; and the prominent actors we have identified above showed a quite regular participation throughout the course. Interestingly the teachers, in spite of their high level of activity in the general workspace, are never on the top position, which is considered a positive indicator in the sense that the students (or at least some of them) indeed got involved in the classroom activities.
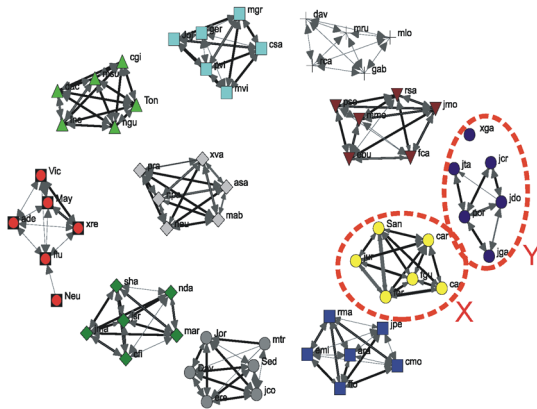

Fig. 2. Sociogram of indirect relationship network at the group workspace.

3.1.2 Study of the Activity at the Private Workspaces

As it has been explained above, the activity at the *private workspaces* has a different nature that the one performed at the general workspace. While the analysis of the general workspace shows the structure that yields from the interactions that take place in the context of discussions that are indirectly related to the fulfillment of the assignments, collaboration at the private group level is mainly focused on the development of a common product (or a set of common products), namely: the set of assignments required by the teacher for the fulfillment of the project tasks.

In order to analyse this workspace, we performed a similar study to the one previously presented. Fig. 2 shows the indirect relationships network at the group private workspaces during the same period of time as the one previously considered. As expected, it displays a rather different structure, due to the restricted access to each group workspace, which results in a set of independent sub-networks that represent the relationships in each group. By simple observation of the sociogram it is easy to detect which were the most and the least active and homogeneous (balanced) groups. These groups are respectively labelled *X* and *Y*. The evolution of these groups is shown by the indexes in Table 2. We can observe that, while group X has a 100% of density and a 0% of centralization during all the phases of the course, group Y never achieves these desired values, with the lowest value at the last period of the course.

This difference is even more outstanding if we compare the valued densities obtained by both groups throughout the course. According to [18], group X shows an ideal pattern of interaction, with all the members interacting with each other, and none of them taking a too central position. Moreover, if we go back now to the indirect relationships at the general workspace network analysed in the previous subsection, we can observe that most of the members of group X are at the center of the network and belong to the list of prominent actors (actually, only *car* does not occupy a central position at the sociogram of Fig. 1). This can mean that the members of group X obtained a good result regarding criteria GF1 and GF2.

Table 2. Evolution of the values of density and centralisation for groups X and Y.

Phase	Group X				Group Y			
	Δ	Δ_v	C_{iD}	C_{oD}	Δ	Δ_v	C_{iD}	C_{oD}
Ph2	100 %	2833 %	0 %	0 %	63,33 %	186,67 %	25 %	25 %
Ph3	100 %	1493 %	0 %	0 %	46,67 %	76,67 %	50 %	50 %
Ph4	100 %	1696 %	0 %	0 %	6,67 %	13,33 %	50 %	20 %
Ph5	100 %	980 %	0 %	0 %	23,33 %	43,33 %	55 %	25 %

However, the fact that a group showed to be effective at this level does not imply that its members show the same effectiveness regarding the acquisition and use of active learning and interaction or processing skills (criteria TP1 and GF3). A tutor or evaluator certainly needs to make use of this important information in order to evaluate each active group member more objectively. To that end, a quantitative (statistical) evaluation technique offers the means to decode the students' active behaviour both at task performance and group functioning level, exploring whether their active participative behaviour at the general and group workspaces also indicates an active and equiponderant contribution to the group's production and well-being functions. Among the groups we analysed statistically, we next chose to discuss the results obtained by the analysis of the interaction behaviour of the members of group X, which presented a particular interest to our research and shows the necessity of applying different evaluation techniques into an integrated approach.

3.2 A Quantitative Statistical Analysis of Task Performance and Group Functioning of an Effective Group

The quantitative analysis performed here is a simple descriptive statistical analysis that aims to provide a complementary and more refined analysis of the contributing and interaction behaviour of the members of an effective group so that to explore and understand the real performance and achievement of each member. This analysis proves to be a necessary aid to the SNA carried out before, especially for identifying particular attitudes of group members not been able to be tracked through SNA.

To illustrate this, we chose to analyse the real effectiveness of group X as concerns criteria TP2 and GF3 (which were partially measured by the qualitative method). On the one hand, the qualitative method assessed this group with an A mark[1] at all evaluation levels; on the other hand, SNA showed it to be one of the most active and

---

[1] We use a 5-point scale mark from A (excellent) to D (fail).

well-balanced groups of the experience whereas its members were also considered as prominent actors (to a greater or lesser degree) in the general classroom workspaces.

The quantitative analysis of criteria TP2 and GF3 depends on the information provided by the data sources available. In our case, BSCW data logs account for various action types performed on several object types.

Consequently, a sufficient measurement of criterion TP2 involves the following generic actions: Create and Change (also called *active learning actions*) as well as the Read action. In particular:

- Create: Document, Note (implies knowledge or information generation that contributes to task performance and group product)
- Change: Document (implies active participation to further elaboration, refinement or revision of existing knowledge)
- Read: Document, Note.

To measure criterion GF3 we can use the following actions which when performed on particular objects enable us to draw conclusions about specific active interaction or processing skills exhibited by each member, such as:

- Task processing skills: can be measured by the actions *create folder,* that indicates task (and indirectly knowledge) management, and *create others* that involves task plannig and preparation of virtual meetings or group agenda).
- Workspace processing skills: can be estimated by the action *move "all" objects* that describes the workspace organisation and maintenance.
- Communication processing skills: can be quantified by the action *change Note/Folder/Other* that adds appropriate meta-information to an object to facilitate and promote better understanding of the objects the members share in the group workspace and thus to achieve more effective interaction.

Table 3 presents a detailed description of the four generic actions and the percentage of objects each member contributed with respect to the other members. For instance, the member *cao* contributed to the creation of the 6.5% of all the documents produced in the group.

The analysis of the data provided in Table 3 allows us to draw a number of conclusions regarding the achievement of criteria TP2 and GF3 by group X.

As for criterion TP2, we observe that there are important differences regarding the contributing behaviour of some members during task realisation; only Read seems to be the most balanced action for all members. In particular, a broad use of active learning skills is shown by member *fer* in generating knowledge (document creation), and by member *fgu* and *fer* as well in generating information (note creation) and document modification (change document). In other words, *fer* and *fgu* show a distinguishing productivity in comparison to their peers as task achievement concerns.

Among the other members, *jur* and *san* have a rather homogeneous, middle-rate contributing activity to the task regarding all actions involved in TP2. In contrast, the use of active learning skills of *cao* and *car* is shown to be the lowest in the group, except *car's* contribution to knowledge generation (had the second highest rate, 15.45%). To interpret *car's* particular attitude to task realisation, we consulted the qualitative evaluation of the individual and group reports, which explained that *car's* focused attention to document creation was primarily due to his particular interest to

specific topics of the case study. In contrast, *cao's* contributing behaviour to knowledge generation deviates significantly from the one of his peers; he also gives the lowest rate at the rest of TP2 actions.

Table 3. Percentage of objects created, read, modified, or moved by each member of group X.

ACTIONS	MEMBERS OBJECTS	Cao	car	fer	fgu	Jur	san
Create	Document	6.5	15.45	38.21	14.63	13.82	11.38
	Note/s	10.05	10.32	20.63	21.69	18.78	18.51
	Folder	17.65	5.88	52.94	5.88	5.88	11.76
	Others	40	26.67	33.33	0	0	0
Read	Document	12.36	16.30	20.90	18.16	15.54	16.74
	Note/s	14.07	14.07	18.98	17.49	17.76	17.61
Change	Document	6.42	8.02	25.67	28.89	13.90	17.11
	Note/s	0	0	20.83	37.50	33.33	8.33
	Folder	0	0	66.67	0	0	33.33
	Others	0	0	100	0	0	0
Move	'all objects'	10.28	4.25	61.54	5.50	7.90	10.50

To measure the achievement of criterion GF3, looking at the use of task processing skills (create folder/others) we observe the distinguished contribution of member *fer* as well as the supporting attitude of the members *cao* and *car*, especially to task planning (Create Others). As for workspace processing skills (*move* action), we notice that member *fer* is the one who cared more for organising and maintaining the group's workspace; a more refined analysis (not shown for lack of space), showed that other members gave only some (very specific) support, like *san* in document organisation, *jur* in message organisation and *cao* in group meetings and agenda organisation. Studying the use of communication processing skills (Change Note/Folder/Other), we again observe the notable contribution of member *fer*, the specific support of other members, like *fgu* and *jur*, to aspects like "communication improvement" (Change note) and member *san* to group workspaces description (Change Folder), and the null contribution of members *cao* and *car*.

As a result of this analysis, group X does not meet criterion GF3 satisfactorily, since its members did not contribute equally to task, workspace and communication processing; indeed, just one member (*fer*) carried the greatest part of the responsibility burden for this particular collaboration feature; the others showed a rather irregular and uneven contribution, supporting only specific aspects as discussed above.

3.3 Discussion: Reflections on the Analyses Performed

Our study showed that evaluation of a real collaborative learning situation is a very complex task, since one has to consider a variety of aspects and thus to integrate several analysis techniques, data and tools into a mixed evaluation method. To that end, after identifying a set of potential indicators of group performance, we first went through a formative qualitative evaluation of all of them. Four of them clearly showed to need a further in-depth analysis to be evaluated completely. The use of two further

different techniques, which proved to be complementary, was guided by the indicators that each technique best accomplished. Classifying each indicator into a specific category (aspect) of the collaborative learning process and assigning it a relative weight dictate the way each technique is used and influences the evaluation process and how it is positioned and related to the others. Exemplifying the evaluation on a specific group showed that the application of different techniques is essential to unfold the group's internal workings and achieve a more objective interpretation of each member attitude and competence.

While the qualitative evaluation and SNA techniques indicated group X as an effective group (excellent group learning outcome, active, well-balanced, with an ideal pattern of interaction and group processing as well as a sufficient social support and help supply behaviour), a further quantitative analysis showed that two specific criteria (TP2 and GF3) were not completely achieved by all members. There were some members who exercised low active learning or interaction/processing skills, like member *cao* who got a low score in specific competences, like knowledge generation, knowledge modification and communication processing, and member *car* whose performance was insufficient regarding knowledge modification and communication processing. In fact, a further refined qualitative study of the individual and group reports confirmed this particular situation.

Indeed, based on an integrated analysis of the interaction of an effective group, we could see that some of its members, who proved to be prominent and influencing at the general workspaces, continued to act at the same level in the group workspace, while others did not act as well at some aspects. In particular, we could deduce that member *fer* kept on distinguishing as a prominent actor both at the general and the group workspaces, yet at all aspects of collaboration. Instead, the activity of members *cao* and *car* at the group space was eclipsed by some other prominent actor (e.g., *fer*). In other words students who played a prominent role at the general workspaces, when put together to collaborate in the same group, managed to develop an excellent group product, however, they did not achieve an ideal synergy at all aspects.

Consequently, it is important for an evaluation method to enable the evaluator to distinguish such particular cases of insufficient contributing or interaction behaviour and allow him/her to infer correct conclusions for the performance and competence of each group member. Doing so, he/she will be able to intervene adequately to monitor and provide the most proper support and guidance when and to whom is needed.

4   Conclusions and Future Work

In this work we have presented an integrated approach to be used by tutors and evaluators of group interaction in order to monitor and assess the performance of virtual learning groups effectively, especially in the case of real, complex and long-term collaborative learning experiences. We demonstrated that particular evaluation techniques tackle specific aspects or indicators of the evaluation process, so an adequate combination of them is needed to cover, complement and verify the gaps or flaws in the analysis results obtained by one technique or another. The conceptualisation and implementation of an integrated approach aims at providing the evaluator flexible options and well-adapted means for building a principled framework that achieves a more complete and effective evaluation of group interaction. We are con-

scious that in order to accomplish an in-depth and objective analysis and evaluation of the entire collaborative learning practice, we need to provide the evaluators efficient, automatic and effective tools to carry out this task within the proposed framework. Given that the theoretical principles and conceptual bases are set and successfully tested by our integrated approach, our objective is now turned to the development of automated tools that will make the tutor's evaluation task much easier and successful, especially as the qualitative and quantitative analysis concerns. As for the former, a couple of tools are currently designed and implemented for the automated processing and analysis of questionnaires and reports related to self-evaluation, group processing as well as to all criteria associated with social support and help supply. As for the latter, the descriptive statistical analysis performed is currently insufficient, since it does not allow us to establish hypothesis that relate the interactions with the individual or group learning outcome, so our aim is to provide a more comprehensive multivariate statistical analysis and build a tool that automates the whole process as much as possible. These tools are to be integrated with SAMSA and with the CSCL tools that provide the interaction data, which makes the definition of clear interfaces between them an interesting line of future research.

Once these new tools and further data sources are available, it would be the right moment to test the framework in a real situation, as the experience is taking place. This can provide the tutor and the group themselves the means that will help them identify the *weaknesses* and the exhibited in a collaborative learning team. As a consequence, this will assist the tutor to understand, monitor, support and assess group interaction and performance in a more efficient way.

Acknowledgments

References

1. Barros, M. & Verdejo, M.: Analysing student interaction processes in order to improve collaboration. The DEGREE approach. Int. J. of Art. Int. in Education. 11 (2000) 221-241
2. Bentley, R. Appelt, W. Busbach. U. Hinrichs, E. Kerr, D. Sikkel, S. Trevor, J. & Woetzel, G.: Basic Support for Cooperative Work on the World Wide Web. Int. J. of Human-Computer Studies 46(6) (1997) 827-846
3. Cho, H. Stefanone, M. & Gay, G.: Social information sharing in a CSCL community. In: G. Stahl (ed.) Computer Support for Collaborative Learning: Foundations for a CSCL community. Erlbaum, NJ (2002) 43-50
4. Collazos, C. Guerrero, L. Pino, J. & Ochoa, S.: Evaluating Collaborative Learning Processes. In: J. M. Haake & J.A. Pino (eds.) Proc. of the 8th Int. Workshop on Groupware (CRIWG 2003). Springer, Berlin (2002) 203-221
5. Daradoumis, T. Guitert, M. Giménez, F. Marquès, J.M. & Lloret, T. : Supporting the Composition of Effective Virtual Groups for Collaborative Learning. In Proc. of the Int. Conf. on Computers in Education (ICCE'02). IEEE Comp. Soc. Press (2002) 332-336
6. Daradoumis T. Xhafa, F. & Marquès J.M.: Evaluating Collaborative Learning Practices in a Virtual Groupware Environment. In: Proc. of the Int. Conf. on Computers and Advanced Technology in Education (CATE 2003), ACTA Press (2003) 438-443

7. Daradoumis T. Xhafa, F. & Marquès J.M.: Exploring Interaction Behaviour and Performance of Online Collaborative Learning Teams. In: Proc. of the 9th Int. Workshop on Groupware (CRIWG 2003), Springer, Berlin (2003) 203-221

8. Dillenbourg, P. (ed.): Collaborative Learning. Cognitive and Computational Approaches. Elsevier Science Ltd. (1999) 1-19

9. Frechtling, J. & Sharp, L. (eds.): User-Friendly Handbook for Mixed Evaluations. Directorate for Education and Human Resources Division of Research, Evaluation and Communication, NSF (1997)

10. MacDonald, J.: Assessing online collaborative learning: process and product. In Int. J. of Computers & Education, 40 (2003) 377-391.

11. McGrath, J.E.: Time, Interaction and Performance (TIP). A Theory of Groups. Small Group Research, 22, (1991) 147-174

12. Martínez, A. Dimitriadis, Y. Rubia, B. Gómez, E. & de la Fuente, P.: Combining qualitative and social network analysis for the study of social aspects of collaborative learning, Computers and Education, 41(4) (2003) 353-368

13. Mühlenbrock, M.: Action-based collaboration analysis for group learning. IOS Press, Amsterdam (2001)

14. Sfard, A.: On two metaphors for learning and the dangers of choosing just one. Educational Researcher 27(2) (1998) 4-13

15. Soller, A. Supporting Social Interaction in an Intelligent Collaborative Learning System. Int. J. of Artificial Intelligence in Education, 12, (2001) 40-62

16. Wasserman, S. & Faust, K.: Social Network Analysis: Methods and Applications. Cambridge Univ. Press, Cambridge (1994)

17. Webb, N.: Testing a theoretical model of student interaction and learning in small groups. In: R. Hertz-Lazarowitz and N. Miller (Eds.), Interaction in Cooperative Groups: The Theoretical Anatomy of Group Learning. Cambridge Univ. Press, NY (1992) 102-119

18. Wortham, D.W.: Nodal and Matrix Analyses of Communication Patterns in Small Groups. Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference. Palo Alto, CA (1999) 681-686